Acta Crystallographica Section F
**Structural Biology and Crystallization Communications**
ISSN 1744-3091

# Crystallization of oligonucleotides containing A-rich repeats suggests a structural contribution to the autoregulation mechanism of PABP translation

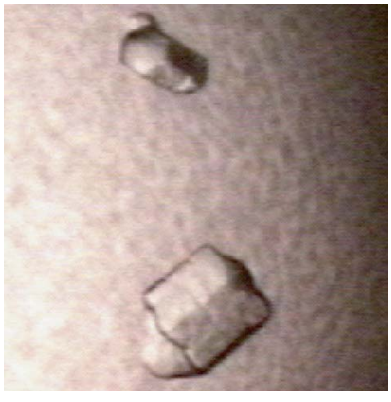
**Keita Kikuchi,[a] Satoru Shimizu,[a] Yoshiteru Sato,[a]‡ Ella Czarina Morishita[a]§ and Akio Takénaka[a,b,c]***

[a]Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Nagatsuda, Midori-ku, Yokohama 226-8501, Japan, [b]Graduate School of Science and Engineering, Iwaki-Meisei University, Chuodai-iino, Iwaki, Fukushima 970-8551, Japan, and [c]Faculty of Pharmacy, Iwaki-Meisei University, Chuodai-iino, Iwaki, Fukushima 970-8551, Japan

‡ Present address: Faculty of Sciences, Kyushu University, 6-10-1 Hakozaki, Higashi-ku, Fukuoka 812-0053, Japan.
§ Present address: Systems and Structural Biology Center, RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan.

Correspondence e-mail: atakenak@iwakimu.ac.jp

Received 12 October 2011
Accepted 2 December 2011





Eukaryotic poly(A)-binding protein (PABP) commonly binds to the 3′-UTR poly(A) tail of every mRNA, but it also binds to the 5′-UTR of PABP mRNA for autoregulation of its expression. In the sequence of the latter binding site, the contiguous A residues are segmented discretely by the insertion of short pyrimidine oligonucleotides as linkers, so that $(A)_{6–8}$ segments are repeated six times. This differs from the poly(A)-tail sequence, which has a higher binding affinity for PABP. In order to examine whether the A-rich repeats have a functional structure, several RNA/DNA analogues were subjected to crystallization. It was found that some of them could be crystallized. Single crystals thus obtained diffracted to 4.1 Å resolution. The fact that the repeated sequences can be crystallized suggests the possibility that the autoregulatory sequence in PABP mRNA has a specific structure which impedes the binding of PABP. When PABP is excessively produced, it could bind to this sequence by releasing the structure in order to interfere with initiation-complex formation for suppression of PABP translation. Otherwise, PABP at low concentration preferentially binds to the poly(A) tail of PABP mRNA.

## 1. Introduction

In protein synthesis in eukaryotes, poly(A)-binding protein (PABP) binds to the poly(A) tail of the 3′-UTR of every mRNA and has several functions in protection of the tail, formation of the initiation complexes of 80S ribosomes and initiation of translation (Svitkin & Sonenberg, 2006; Kahvejian *et al.*, 2005; Karim *et al.*, 2006; Mangus *et al.*, 2003; Kozlov *et al.*, 2001; Fig. 1*a*). Interestingly, it has been found that PABP also binds to the 5′-UTR of PABP mRNA with a binding affinity that is lower than that to the poly(A) tail. This has been interpreted as the suppression of translation of PABP on its overexpression (Grange *et al.*, 1987; de Melo Neto *et al.*, 1995; Bag & Wu, 1996; Wu & Bag, 1998; Bag, 2001), as shown in Fig. 1(*b*).

The nucleotide sequences of the two binding sites are definitely different. The 3′-UTR site has an extended poly-A sequence that is composed of contiguous A residues. In contrast, the 5′-UTR site contains an A-rich sequence in which the contiguous A residues are discretely segmented into short lengths containing 6–8 A residues each and the segments are joined together by short linkers of 3–5 U/C residues (Grange *et al.*, 1987), as shown in Fig. 1(*c*). This difference suggests that the binding mechanism of PABP to the 5′-UTR is different from that to the 3′-UTR poly(A) tails of every mRNA, with the latter being proposed from the crystal structure of PABP in complex with an oligo-A (Deo *et al.*, 1999). There is a question as to why it is necessary for the A residues in the 5′-UTR to be segmented into short lengths separated by spacers or gaps. It is speculated that the two distinct sequences could be designed to adopt different secondary and tertiary structures in order to control PABP binding with different binding affinities. There are examples in which the 5′-UTR structure acts as an enhancer of mRNA translation (Vivinus *et al.*, 2001) or facilitates efficient protein-synthesis initiation (Hess & Duncan, 1996). In order to examine whether the 5′-UTR A-rich autoregulatory sequence can have a specific structure prior to and

**Table 1**
RNA and DNA fragments prepared in this study.

x represents $C^{Br}$ (a 5-bromocytosine residue) and A, c, u and t represent adenine, cytosine, uracil and thymine residues, respectively.

| Name | Residues | Repeats | Sequence |
|---|---|---|---|
| RNA oligomers | | | |
| 16-mer | 16 | 2 | 5′-AAAAAAuxuAAAAAAc-3′ |
| 25-mer | 25 | 3 | 5′-AAAAAAuxuAAAAAAucuAAAAAAc-3′ |
| 34-mer | 34 | 4 | 5′-AAAAAAuxu(AAAAAAucu)$_2$AAAAAAc-3′ |
| DNA oligomers | | | |
| 7-mer | 7 | 1 | 5′-AAAAAAc-3′ |
| 16-mer | 16 | 2 | 5′-AAAAAAtctAAAAAAc-3′ |
| 25-mer | 25 | 3 | 5′-(AAAAAAtct)$_2$AAAAAAc-3′ |
| 34-mer | 34 | 4 | 5′-(AAAAAAtct)$_3$AAAAAAc-3′ |
| 45-mer | 45 | 5 | 5′-t(AAAAAAtct)$_4$AAAAAAtc-3′ |

after PABP binding and, if so, to determine the structure and to elucidate the role of such a structure in the PABP autoregulation mechanism, we have surveyed the possibility of crystallizing several RNA and DNA fragments containing regular A-rich repeats.

## 2. Materials and methods

The main purpose of the present study is to examine the possibility that A-rich repeats can form a structure. For this reason, the native sequence, which contains irregular parts, was modified for crystallization. The oligonucleotides listed in Table 1 were synthesized using a DNA/RNA synthesizer. In a similar manner, several deoxy-oligonucleotides (Table 1) were also synthesized to compensate for the difficulty of RNA crystallization. These oligomers, which were pre-purified by gel filtration on a Sephadex G-25 (Amersham) column, were further purified by electrophoresis on large-size (400 × 300 mm) gels. The gel regions containing the samples were cut out by monitoring the fluorescent bands. The eluted samples were finally purified using a combination of SuperQ 650C (Toyopearl) ion-exchange and long Sephadex G-25 columns and were then dried in a frozen state. The purities of the samples were confirmed by SDS–PAGE. Calculated absorption coefficients of the samples were used to estimate their concentrations for crystallization.

Crystallization was examined using the hanging-drop vapour-diffusion method at 278 and 310 K with several crystallization kits from Hampton Research. The preliminary conditions under which crystalline deposits appeared were further optimized by varying the temperature and the concentration of the sample, the buffer pH and the precipitants. Microscopy with fluorescence spectroscopy, which was devised for nucleic acid crystallization (Kondo & Takénaka, 2005), was used to judge whether the crystals contained nucleic acids. Crystals that were suitable for X-ray experiments were mounted in nylon cryoloops (Hampton Research) with the mother liquor containing 30%(*v*/*v*) MPD as a cryoprotectant and were stored in liquid nitrogen prior to X-ray experiments. X-ray data were obtained at 100 K on the AR-NW12 beamline of the Photon Factory (PF) synchrotron-radiation facility (Tsukuba, Japan). Diffraction patterns were processed with the programs *HKL*-2000 (Otwinowski & Minor, 1997) and *CrystalClear* (Pflugrath, 1999).

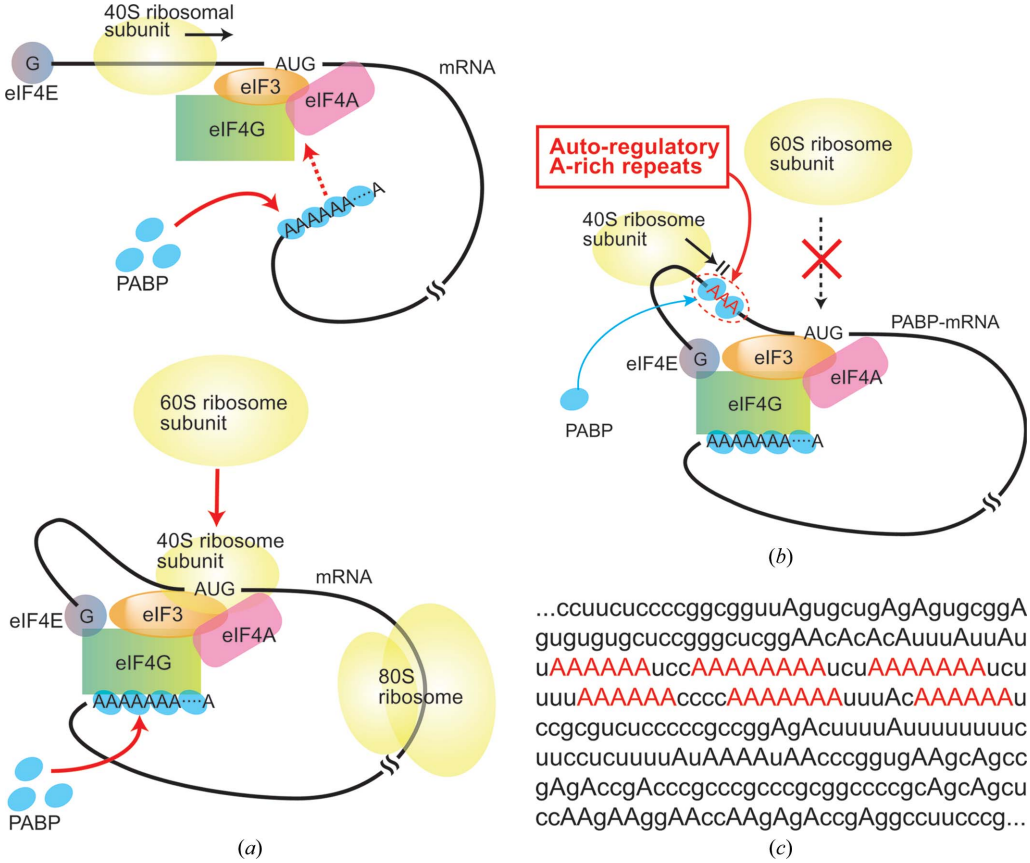



**Figure 1**
Proposed roles of PABP (*a*) in the initiation-complex formation for protein synthesis and (*b*) in the autoregulation mechanism for PABP suppression; (*c*) the autoregulatory sequence found in the 5′-UTR of human PABP mRNA. eIF is a eukaryote initiation factor, eIF3 and eIF4A are 40S ribosome-associated initiation factors, eIF4G is a scaffolding protein that promotes formation of the closed loop of mRNA and eIF4E is a cap-binding protein.

## 3. Results and discussion

Fig. 2 shows that the synthesized samples are purified to a sufficient level for general crystallization. The 34-mer and 45-mer DNAs have successfully been crystallized (Fig. 3), while some of the RNAs were obtained as crystalline precipitates. The conditions are given in Table 2. The 34-mer crystals appeared to be single and their fluorescent microscopic images showed that they are composed of nucleic acids. The largest crystal grew to dimensions of 0.3 × 0.2 × 0.1 mm within 11 d. One of the X-ray diffraction patterns is shown in Fig. 4; the crystal diffracted to a maximum resolution of 4.1 Å. After indexing of diffraction spots, the unit-cell parameters were estimated to be $a = b = 46.1$, $c = 229.1$ Å in a hexagonal primitive lattice. The 45-mer crystals also appeared to be single, but are presently too small for X-ray experiments.

In the present study, we have found that both the DNA and RNA fragments containing tandem AAAAAA repeats are crystallizable

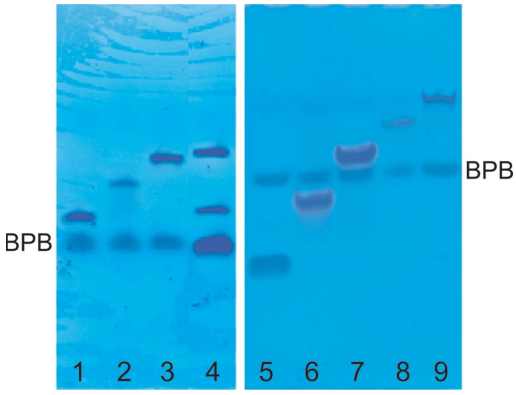



**Figure 2**
Purity assay by SDS–PAGE of the synthesized samples; lanes 1, 2 and 3 contain 16-mer, 25-mer and 34-mer RNAs, respectively; lanes 5, 6, 7, 8 and 9 contain 7-mer, 16-mer, 25-mer, 34-mer and 45-mer DNAs, respectively. Nucleotide bands were stained with toluidine blue. Electrophoresis was monitored with bands of xylene cyanol and bromophenol blue (labelled BPB). The marker in lane 4 contains (from the bottom) 10-mer, 20-mer and 40-mer oligonucleotides. The 10-mer band is strengthened by overlap with the bromophenol blue band.

**Table 2**
Crystallization conditions.

MPD, 2-methyl-2,4-pentanediol; MgAc, magnesium acetate tetrahydrate.

| | 34-mer RNA | | 34-mer DNA | | |
|---|---|---|---|---|---|
| Crystal | Form 1 | Form 2 | Form 1 | Form 2 | 45-mer DNA |
| Droplet | | | | | |
| Sample (m*M*) | 0.25 | 0.25 | 0.5 | 0.5 | 0.5 |
| Na cacodylate (m*M*) | 25 (pH 6.5) | 25 (pH 6.5) | 20 (pH 6.0) | 40 (pH 6.0) | 25 (pH 7.0) |
| Spermine.4HCl (m*M*) | 6 | 6 | 10 | 13 | 5 |
| NaCl (m*M*) | — | — | 6 | 50 | 50 |
| KCl (m*M*) | — | — | 60 | 50 | — |
| $Mg^{2+}$ (m*M*) | 5 ($MgSO_4$) | 5 (MgAc) | — | — | 50 ($MgCl_2$) |
| $(NH_4)_2SO_4$ (*M*) | 1 | — | — | — | — |
| $Li_2SO_4{\cdot}H_2O$ (*M*) | — | 0.7 | — | — | — |
| MPD (%) | — | — | 5 | 2 | 5 |
| Reservoir | | | | | |
| MPD (%) | — | — | 20 (1 ml) | 8 (1 ml) | 30 (1 ml) |
| $(NH_4)_2SO_4$ (*M*) | 2 (0.7 ml) | — | — | — | — |
| $Li_2SO_4{\cdot}H_2O$ (*M*) | — | 1.3 (0.7 ml) | — | — | — |
| Temperature (K) | 310 | 310 | 277 | 277 | 277 |

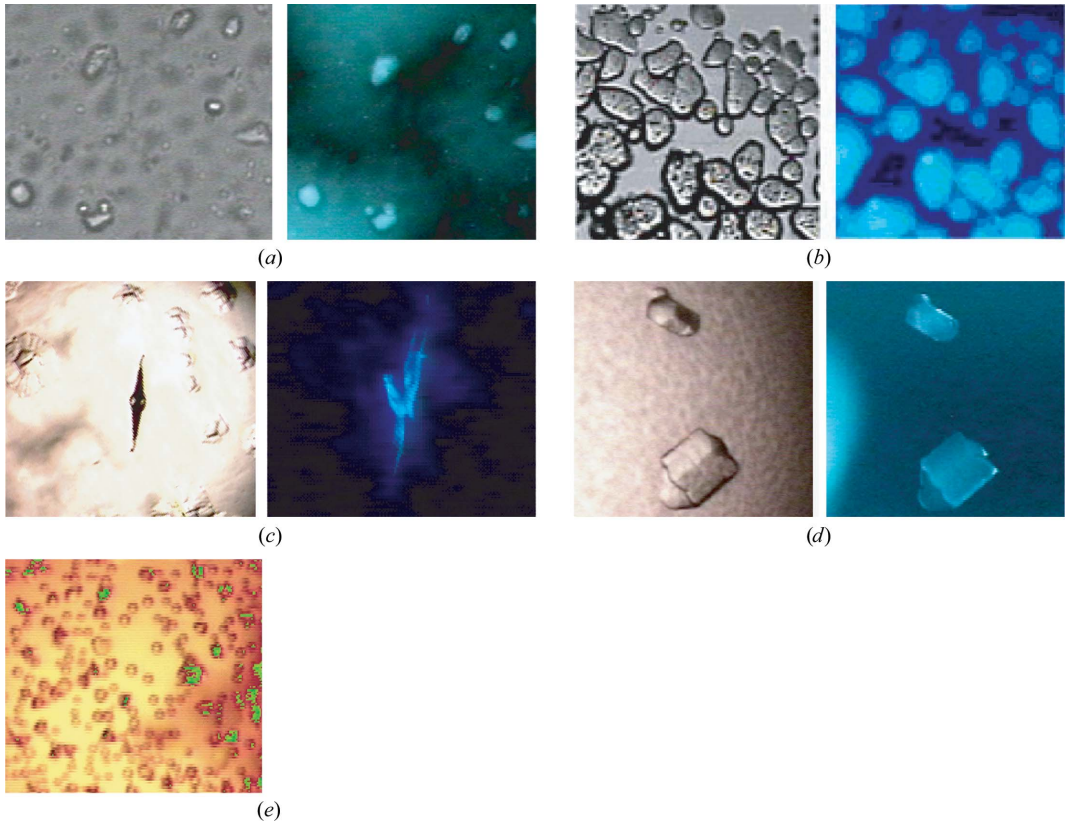



**Figure 3**
Crystals of A-rich repeats obtained under the conditions given in Table 2: (*a*) 34-mer RNA (form 1), (*b*) 34-mer RNA (form 2), (*c*) 34-mer DNA (form 1), (*d*) 34-mer DNA (form 2) and (*e*) 45-mer DNA. The right-hand pictures in (*a*–*d*) show the corresponding fluorescent images.

and that the RNA fragments can also be crystallized. These facts suggest that AAAAAA repeats adopt a regular structure in the cases of both DNA and RNA. Therefore, it could be speculated that the autoregulatory sequence in PABP mRNA has a specific structure that facilitates translation. When PABP is excessively produced, it will bind to this sequence to release the structure so that formation of the initiation complex will be prevented in order to suppress PABP translation. Otherwise, at a low concentration of PABP the protein can easily bind to the poly(A) tails. Such a preference occurs in the nonstructural poly(A) part because it has no regularity to maintain the specific structure, even if the adenine residues interact with each other in the tails.

It is of interest to discuss possible structures of DNA and RNA crystals containing A-rich repeats similar to the autoregulatory sequence. The structural features will be formed through hydrogen bonds between the A-rich repeats. Three basic motifs may possibly occur for A–A interactions: motif 1 is formed between the Watson–Crick edges through $N^6$—H···$N^1$ hydrogen bonds, motif 2 occurs between the Hoogsteen edges through $N^6$—H···$N^7$ hydrogen bonds and motif 3 occurs between the Watson–Crick edge and the Hoogsteen edge, as shown in Fig. 5(*a*). When the autoregulatory sequence is folded at the linker regions, as shown in Fig. 5(*b*), the stacked adenine bases could form hydrogen bonds to each other between the columns. In combination with motifs I and II or with motif III, the adenine columns can be assembled to form an A-rich cluster (Cai *et al.*, 1998). To confirm the formation of an A-rich cluster, it is necessary to obtain crystals which diffract to higher resolution. In addition, it is of interest to reveal the structure of A-rich repeats in complex not only with PABP but also with PABP and PABP-associated proteins, because this autoregulatory sequence is considered to be involved in the formation of a heteromeric ribonucleoprotein complex *in vivo* (Patel *et al.*, 2005).

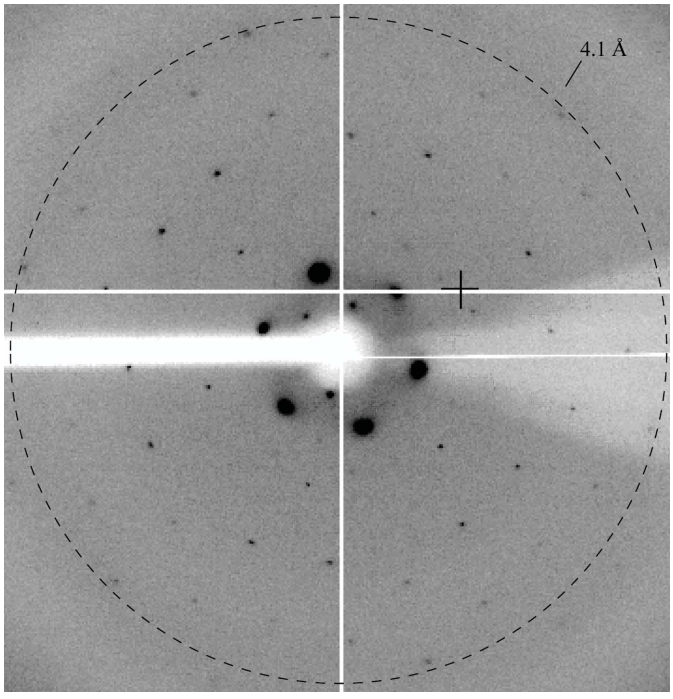



**Figure 4**
An X-ray diffraction pattern of 34-mer DNA (form 2) containing four AAAAAA repeats obtained on beamline AR-NW12 at PF with a wavelength of 1.00 Å, an exposure time of 5 s, a crystal-to-detector distance of 250 mm and a temperature of 100 K.

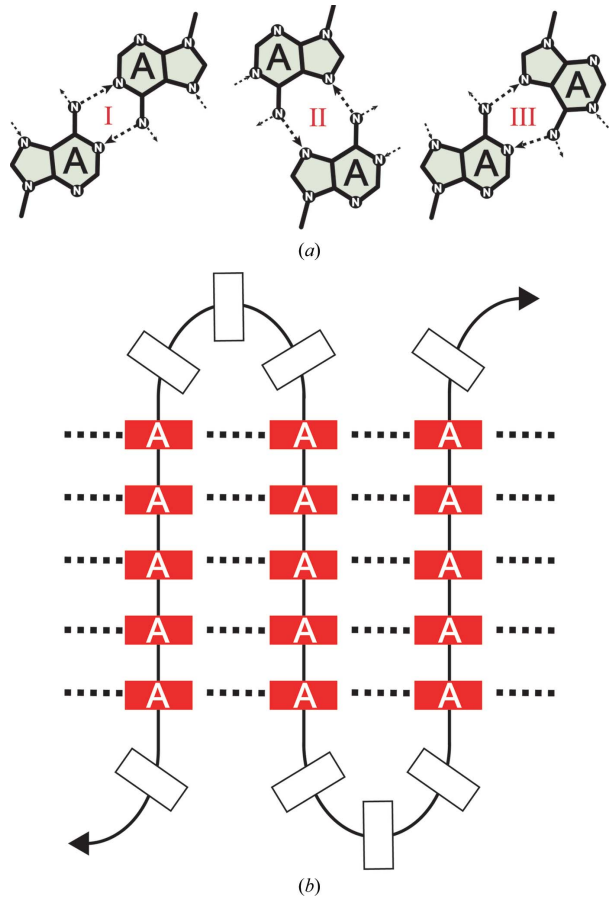



**Figure 5**
(*a*) Three possible hydrogen-bonding motifs between the two adenine bases. The A bases in a segment can stack on each other to form a column. The stacked columns can interact with each other through the two hydrogen bonds in motif III to form a cluster, as shown in (*b*). An alternative combination of motifs I and II also makes a cluster of the columns. The arrows in (*a*) indicate the direction of a hydrogen bond from donor to acceptor.

This work was supported in part by a Grant-in-Aid for a research program in RNA structural biology from the Ministry of Education, Culture, Sports, Science and Technology of Japan. We thank G. Kawai for organizing the research group in the program and N. Igarashi and S. Wakatsuki for facilities and help during data collection.

## References

Bag, J. (2001). *J. Biol. Chem.* **276**, 47352–47360.
Bag, J. & Wu, J. (1996). *Eur. J. Biochem.* **237**, 143–152.
Cai, L., Chen, L., Raghavan, S., Ratliff, R., Moyzis, R. & Rich, A. (1998). *Nucleic Acids Res.* **26**, 4696–4705.
Deo, R. C., Bonanno, J. B., Sonenberg, N. & Burley, S. K. (1999). *Cell*, **98**, 835–845.
Grange, T., de Sa, C. M., Oddos, J. & Pictet, R. (1987). *Nucleic Acids Res.* **15**, 4771–4787.
Hess, M. A. & Duncan, R. F. (1996). *Nucleic Acids Res.* **24**, 2441–2449.
Kahvejian, A., Svitkin, Y. V., Sukarieh, R., M'Boutchou, M. N. & Sonenberg, N. (2005). *Genes Dev.* **19**, 104–113.
Karim, M. M., Svitkin, Y. V., Kahvejian, A., De Crescenzo, G., Costa-Mattioli, M. & Sonenberg, N. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 9494–9499.
Kondo, J. & Takénaka, A. (2005). *Protein Crystallization for Diffraction Structural Biology*, pp. 132–146. Kyoto: Kyoto University Press.
Kozlov, G., Trempe, J. F., Khaleghpour, K., Kahvejian, A., Ekiel, I. & Gehring, K. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 4409–4413.
Mangus, D. A., Evans, M. C. & Jacobson, A. (2003). *Genome Biol.* **4**, 223.

Melo Neto, O. P. de, Standart, N. & Martins de Sa, C. (1995). *Nucleic Acids Res.* **23**, 2198–2205.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Patel, G. P., Ma, S. & Bag, J. (2005). *Nucleic Acids Res.* **33**, 7074–7089.
Pflugrath, J. W. (1999). *Acta Cryst.* D**55**, 1718–1725.
Svitkin, Y. V. & Sonenberg, N. (2006). *Mol. Biol.* **40**, 611–619.
Vivinus, S., Baulande, S., van Zanten, M., Campbell, F., Topley, P., Ellis, J. H., Dessen, P. & Coste, H. (2001). *Eur. J. Biochem.* **268**, 1908–1917.
Wu, J. & Bag, J. (1998). *J. Biol. Chem.* **273**, 34535–34542.